

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 04-273299

(43)Date of publication of application : 29.09.1992

(51)Int.Cl.

G10L 3/00

(21)Application number : 03-033967

(71)Applicant : RICOH CO LTD

(22)Date of filing : 28.02.1991

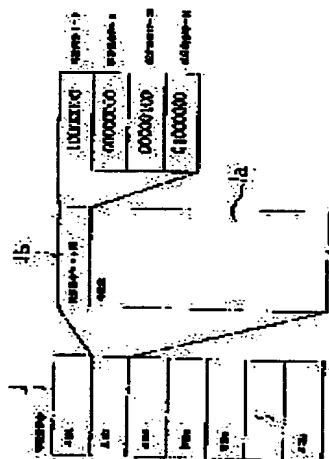
(72)Inventor : MIYAMOTO KEIICHI

(54) VOICE RECOGNITION DEVICE

(57)Abstract:

PURPOSE: To accurately select a partial set dictionary without doubling the word dictionary nor selecting an unnecessary word as an object of recognition even when the same word are used in different situations.

CONSTITUTION: The word dictionary 1 has a bit array 1b corresponding to each partial set in each word in the dictionary and the bit array 1b is turned on and off to indicate which partial set dictionary the word belongs to; and a result selection part outputs information showing which bit array a partial set dictionary to be recognized next corresponds to a partial dictionary selection part to select the object of recognition.



引用文献 2

(19) 日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平4-273299

(43) 公開日 平成4年(1992)9月29日

(51) Int.Cl.⁵
G 10 L 3/00

識別記号 庁内整理番号
301 C 8842-5H
D 8842-5H

F I

技術表示箇所

審査請求 未請求 請求項の数 3 (全 5 頁)

| | |
|-----------|-----------------|
| (21) 出願番号 | 特願平3-33967 |
| (22) 出願日 | 平成3年(1991)2月28日 |

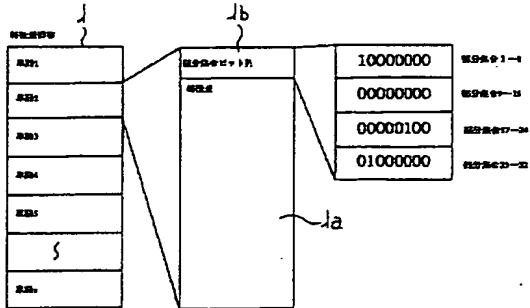
| | |
|----------|--|
| (71) 出願人 | 000006747 株式会社リコー 東京都大田区中馬込1丁目3番6号 |
| (72) 発明者 | 宮本 恵一 東京都大田区中馬込1丁目3番6号 株式 会社リコー内 |
| (74) 代理人 | 弁理士 柏木 明 |

(54) 【発明の名称】 音声認識装置

(57) 【要約】

【目的】 同じ単語が異なる局面で使用されるような場合であっても、単語辞書を二重化したり余分な単語を認識対象としてしまうことなく、部分集合辞書を的確に選定できるようにする。

【構成】 単語辞書1が辞書中の各単語に各部分集合に対応するビット列1bを持ち、そのビット列1bのオン・オフによりその単語が何れの部分集合辞書に属するかを表すものとし、部分辞書選択部に対して結果選択部より次回の認識対象の部分集合辞書がビット列の何れに対応しているかの情報を出力して、認識対象を選定するようにした。



1

2

【特許請求の範囲】

【請求項1】 各部分集合に対応し各ビットのオン・オフによりその単語が何れの部分集合辞書に属するか否かを決定するビット列を各単語毎に有する単語辞書と、入力された単語音声の特徴量を抽出する特徴抽出部と、入力された単語音声の特徴量と予め前記単語辞書に記憶された複数の単語の特徴量との類似度を計算する類似度計算部と、計算された類似度から單一又は複数の単語を認識結果として選択するとともに次回の認識対象とする部分集合辞書が前記単語辞書中の各単語が持つ前記ビット列の何れに対応するかの情報を outputする結果選択部と、この結果選択部による認識結果とビット列情報とに基づき前記単語辞書中の單一又は複数の単語を部分集合として選択して次回の音声認識時にこの選択された部分集合辞書のみを認識対象として選択する部分辞書選択部とよりなることを特徴とする音声認識装置。

【請求項2】 部分辞書選択部に対して結果選択部からの認識結果とともに外部からの部分辞書指定情報を入力させたことを特徴とする請求項1記載の音声認識装置。

【請求項3】 単語辞書が、複数の部分集合辞書に属する単語に対してはビット列情報と複数の部分集合辞書に属することを示す識別子を持ち、ただ一つの部分集合辞書にのみ属する単語に対してはその部分集合を表す識別子のみを持ち、全ての部分集合辞書に属する単語に対しては全ての場合に認識対象となる旨を表す識別子を持つことを特徴とする請求項1又は2記載の音声認識装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、音声認識装置に関する。

【0002】

【従来の技術】 近年、音声認識装置は各種制御機器、パソコン用コンピュータ、ワードプロセッサ等の入力装置として実用化されつつある。これらの装置の認識対象は、一般には、離散発声された単語で制御コマンドや制御オブジェクトを表すものが主である。これらの音声認識・入力装置にとって最も重要な性能は、認識率と認識速度である。

【0003】 ここに、音声入力装置の認識率はその認識対象単語の数の増加に伴い低下し、認識速度も低下する。このため、認識対象単語数はある程度以下に抑えるのが望ましいが、これは音声処理部が勝手に決め得る要素ではない。

【0004】 ところが、コマンドやオブジェクトの構造が階層的となっているとき、例えば、住所を音声で入力するような場合（東京、江東区、××町、○○番地、…等）には、各々の局面で認識対象となる単語は限られている。つまり、都道府県を入力する局面では、市区町村を示す単語は必要ないということなどである。

【0005】 このようなアプリケーションの場合には、

10

音声認識の対象となる単語を、各々の局面で必要な単語だけに限定することが可能である。こうすれば、その局面において、認識対象単語の数が減少したことになり、認識率と認識速度の低下を最低限に抑えることが可能となる。

【0006】 そこで、従来にあってはこのような目的を達成するために認識単語辞書をカテゴリー毎に幾つかに分割し、認識処理時には前もってどの部分辞書を使うかを指定するという方法がとられている。より具体的には、各単語に辞書中の単語属性として、所属する部分辞書のID番号を登録しておき、認識時に指定されるIDがそれと合致するときだけ、認識対象になるというものである。図7はこのようなID番号による単語辞書の分割の様子を示すものである。

【0007】

【発明が解決しようとする課題】 ところが、このような方法では、一つの単語は一つのID番号しか持てないため、同じ単語が異なった局面で使われるときには、若干の工夫が必要となる。例えば、同じ単語辞書を複数作っておき、互いに異なるID番号を与えておくとか、認識時に必要な単語が全て得られるよう複数のID番号を指定するといった方法である。ところが、前者の場合には辞書容量が増大してしまい、後者の場合には複数のID番号を指定するために余分な単語まで認識対象してしまうおそれがあり、当初の目的に反するものとなってしまう。

【0008】

【課題を解決するための手段】 請求項1記載の発明では、各部分集合に対応し各ビットのオン・オフによりその単語が何れの部分集合辞書に属するか否かを決定するビット列を各単語毎に有する単語辞書と、入力された単語音声の特徴量を抽出する特徴抽出部と、入力された単語音声の特徴量と予め前記単語辞書に記憶された複数の単語の特徴量との類似度を計算する類似度計算部と、計算された類似度から單一又は複数の単語を認識結果として選択するとともに次回の認識対象とする部分集合辞書が前記単語辞書中の各単語が持つ前記ビット列の何れに対応するかの情報をoutputする結果選択部と、この結果選択部による認識結果とビット列情報とに基づき前記単語辞書中の單一又は複数の単語を部分集合として選択して次回の音声認識時にこの選択された部分集合辞書のみを認識対象として選択する部分辞書選択部とを設けた。

【0009】 請求項2記載の発明では、部分辞書選択部に対して結果選択部からの認識結果とともに外部からの部分辞書指定情報を入力させた。

【0010】 さらに、請求項3記載の発明では、単語辞書が、複数の部分集合辞書に属する単語に対してはビット列情報と複数の部分集合辞書に属することを示す識別子を持ち、ただ一つの部分集合辞書にのみ属する単語に

50 対してはその部分集合を表す識別子のみを持ち、全ての

3

部分集合辞書に属する単語に対しては全ての場合に認識対象となる旨を表す識別子を持つものとした。

【0011】

【作用】請求項1記載の発明によれば、単語辞書が辞書中の各単語に各部分集合に対応するビット列を持ち、そのビット列のオン・オフによりその単語が何れの部分集合辞書に属するかを表すものとし、部分辞書選択部に対して結果選択部より次回の認識対象の部分集合辞書がビット列の何れに対応しているかの情報を出力して、認識対象を絞るので、局面により認識対象となる単語の重複があつても、単語辞書を二重化したり、余分な単語を認識対象としてしまうことなく、ビット列で表現し得る分の部分集合辞書に対応できるものとなる。

【0012】この際、請求項2記載の発明によれば、部分辞書選択部に対して外部からの部分辞書指定情報の入力も可能としたので、部分集合辞書情報の初期化等に活用できる。

【0013】また、請求項3記載の発明によれば、単語の部分集合辞書に属する場合を3種に分け、その種別を表す識別子を用いて省略された形でビット列情報を単語辞書に格納するようにしたので、単語辞書の容量を大幅に節約できるものとなる。

【0014】

【実施例】本発明の第一の実施例を図1及び図2に基づいて説明する。図2は全体構成を示すブロック図である。まず、予め複数単語分の特徴量を記憶して単語辞書となる特徴量辞書1が設けられている。また、マイクロフォン等を通じて入力される音声入力からその特徴量を抽出する特徴抽出部2が設けられている。この特徴抽出部2と前記特徴量辞書1とは類似度計算部3が接続されている。この類似度計算部3の出力側には結果選択部4が接続され、この結果選択部4と前記特徴量辞書1との間には部分辞書選択部5が接続されている。

【0015】このような構成において、基本的には、音声入力は特徴抽出部2で特徴量が抽出される。抽出された特徴量は類似度計算部3において、部分辞書選択部5によって特徴量辞書1中から選択された語彙のみと類似度が計算される。全ての選択された語彙との類似度が計算された後、最も類似度が高い語が認識結果として選択される。この認識結果は、次回の部分辞書の選択にも使われる。

【0016】ここに、本実施例の特徴とする特徴量辞書1の構造を図1に示し、その辞書選択法を説明する。まず、この特徴量辞書1は記憶された単語数n分の辞書に分割できる。さらに、各単語の特徴量の格納部分1aの一部、例えば図示のように先頭部分には部分集合辞書に対応させた部分集合ビット列の格納部分1bが配置されている。本例では、部分集合1~8, 9~16, 17~24, 25~32で示すように32ビットのビット列が用いられている。このビット列により、各単語は、32

10 4

通りの部分集合辞書に独立して属したり属さなかつたりすることができる。属すか属さないかは、対応するビットのオン・オフ(1又は0)により決まる。図示例は、単語2が部分集合辞書の番号1, 22, 26に属していることを表している。

【0017】このような特徴量辞書1構造に対して、結果選択部4では前回の認識結果や制御される機器の状態に従って、次回の認識に必要な辞書を選択し、その番号を部分辞書選択部5に送る。部分辞書選択部5ではその辞書番号に対応するビットがオンのビット列を有する単語だけを類似度計算部3に送り、認識対象とする。このようにして、各局面に最適な部分集合辞書が選択されて認識が行なわれる。

【0018】なお、音声認識のアルゴリズムや、特定話者・不特定話者方式等は特に限定されるものではなく、また、周知の認識処理部分については説明を省略した。

【0019】つづいて、本発明の第二の実施例を図3ないし図6により説明する。前記実施例で示した部分と同一部分は同一符号を用いて示す。図3に示す全体構成においては、前記実施例とほぼ同様であるが、本実施例では、部分辞書選択部5に対して、結果選択部4からの認識結果の他、部分辞書指定入力を外部から与え得るように構成したものである。このような外部入力を可能とすることにより、初期化などに活用できる。

【0020】ついで、本実施例による特徴量辞書1の構造を説明する。この場合も、記憶された単語数n分の辞書に分割し得る他、各単語の特徴量の一部、例えば先頭部分に、部分集合辞書に対する種別を表す識別子と、種別に応じた部分集合辞書に対応するビット列、辞書番号又はビット列なしとの組合せが格納されている。

【0021】まず、図4は例えばある単語2が複数の部分集合辞書に属する場合を示す。これは、基本的には、図1に示したものと同様であるが、部分集合ビット列情報において、その先頭部分に複数の部分集合辞書に属する旨の識別子“00”を有するものとされている。

【0022】図5は例えばある単語3がただ一つの部分集合辞書に属する場合を示す。この場合には、部分集合ビット列情報としては、この旨を表す識別子“10”と、上記のビット列に代えて、部分集合辞書番号の情報が記憶されている。この辞書番号は、ビットを数値でエンコードした形で記述できるので、32種の部分集合辞書から一つを指定するのに5ビットで足りる。図示例は、“11010”により単語3が部分集合辞書番号26に属していることを示す。

【0023】図6は例えばある単語4が全ての部分集合辞書に属する場合を示す。この場合には、部分集合ビット列情報としては、この旨を表す識別子“01”的みが記憶され、これ以外は記憶されない。図示例は、識別子“01”により単語4が全ての場合に認識対象となることを示す。

5

【0024】辞書選択処理は、前記実施例の場合と同様に行なわれる。

【0025】本実施例によれば、前記実施例に加えて、各単語が部分集合辞書に属する態様によって3種に分けられた辞書構成としたので、部分集合ビット列情報格納による辞書容量の増加を抑制できるものとなる。

【0026】

【発明の効果】本発明は、上述したように構成したので、請求項1記載の発明によれば、単語辞書が辞書中の各単語に各部分集合に対応するビット列を持ち、そのビット列のオン・オフによりその単語が何れの部分集合辞書に属するかを表すものとし、部分辞書選択部に対して結果選択部より次回の認識対象の部分集合辞書がビット列の何れに対応しているかの情報を出力させて、認識対象を絞るようにしたので、局面により認識対象となる単語の重複があっても、単語辞書を二重化したり、余分な単語を認識対象としてしまうようなことなく、ビット列で表現し得る分の部分集合辞書に対処することができる。

【0027】この際、請求項2記載の発明によれば、部分辞書選択部に対して外部からの部分辞書指定情報の入力も可能としたので、部分集合辞書情報の初期化等に活用することができる。

6

【0028】また、請求項3記載の発明によれば、単語の部分集合辞書に属する場合を3種に分け、その種別を表す識別子を用いて省略された形でビット列情報を単語辞書に格納するようにしたので、単語辞書の容量を大幅に節約することができる。

【図面の簡単な説明】

【図1】本発明の第一の実施例を示す特微量辞書の構造図である。

【図2】全体的構成を示すブロック図である。

【図3】本発明の第二の実施例を示すブロック図である。

【図4】特微量辞書の構造図である。

【図5】特微量辞書の構造図である。

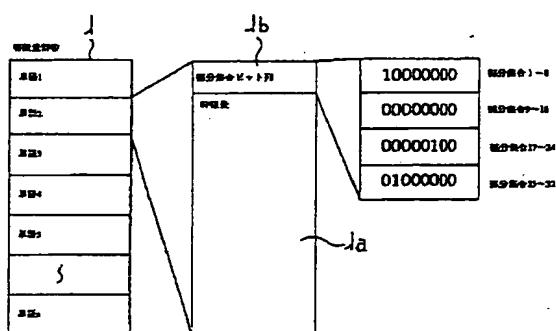
【図6】特微量辞書の構造図である。

【図7】従来例を示す辞書の構造図である。

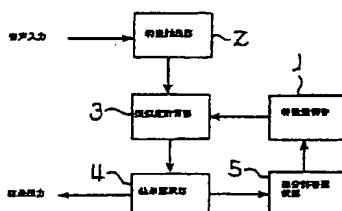
【符号の説明】

- | | |
|-----|---------|
| 1 | 単語辞書 |
| 1 b | ビット列 |
| 2 | 特微量抽出部 |
| 3 | 類似度計算部 |
| 4 | 結果選択部 |
| 5 | 部分辞書選択部 |

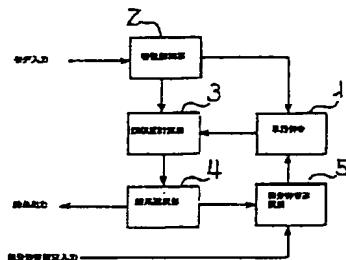
【図1】



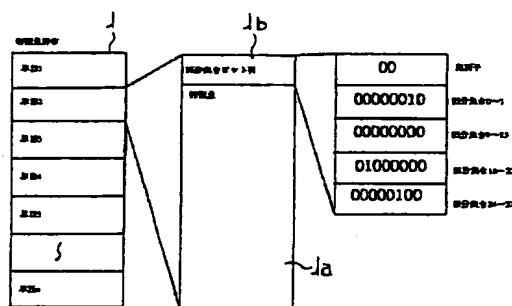
【図2】



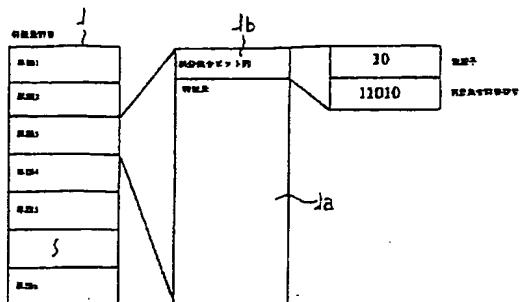
【図3】



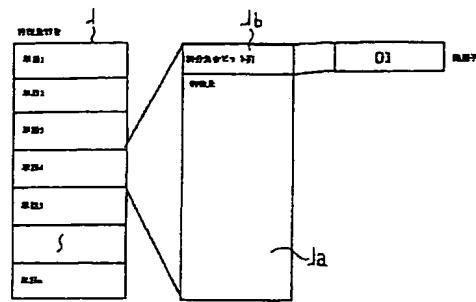
【図4】



【図5】



【図6】



【図7】

